

gaMYB1_alignment_and_sequence.txt

Exhibit 2

gi|7438345|pir||T09879 myb-related protein A - upland cotton
gi|437327|gb|AAA33067.1| MYB1 [Gossypium hirsutum]
gi|23476277|gb|AAN28269.1| myb-like transcription factor 1 [Gossypium hirsutum]

Length = 294
Score = 602 bits (1553), Expect = e-171
Identities = 288/294 (97%), Positives = 291/294 (98%)
Frame = +3

Query: 144 MGRSPCCEKAHTNKGAWTKEEDQRLINIRVHGEWCWRS LPKAAGLLRCGKSCRLRWINY 323
MGRSPCCEKAHTNKGAWTKEEDQRLINIRVHGEWCWRS LPKAAGLLRCGKSCRLRWINY
Sbjct: 1 MGRSPCCEKAHTNKGAWTKEEDQRLINIRVHGEWCWRS LPKAAGLLRCGKSCRLRWINY 60
Query: 324 LRPDLKRGNFTEEEDELI IKLHSL LGNKWSLIAGRLPGRTDNEIKNYWNTHIKRKLISRG 503
LRPDLKRGNFTEEEDELI IKLHSL LGNKWSLIAGRLPGRTDNEIKNYWNTHIKRKLISRG
Sbjct: 61 LRPDLKRGNFTEEEDELI IKLHSL LGNKWSLIAGRLPGRTDNEIKNYWNTHIKRKLISRG 120
Query: 504 IDPQTHRPLNQTAITNTVTGPTELDFRNSPTSVSKSSSIKNPSLDFNYNEFHKSHTDSL 683
IDPQTHRPLNQTA TNTVT PTELD FRNSPTSVSKSSSIKNPSLDFNYNEF FKS+TDSL
Sbjct: 121 IDPQTHRPLNQTAITNTVTAPTELD FRNSPTSVSKSSSIKNPSLDFNYNEFQFKSNTDSL 180
Query: 684 EEPNCTASTGMTTDEEQEQLHKKQYGPSNGQDINLELSIGIVSADSSRVSSANSAESK 863
EEPNECTAS+GMTTDEEQEQLHKKQYGPSNGQDINLELSIGIVSADSSRV S+ANSAESK
Sbjct: 181 EEPNCTASSGMTTDEEQEQLHKKQYGPSNGQDINLELSIGIVSADSSRVSNANSAESK 240
Query: 864 PKVDNNNFQFLEQAMVAKAVCLCWQLGFGTSEICRNCQNSNSNGFYCYRPLDS 1025
PKVDNNNFQFLEQAMVAKAVCLCWQLGFGTSEICRNCQNSNSNGFYCYRPLDS
Sbjct: 241 PKVDNNNFQFLEQAMVAKAVCLCWQLGFGTSEICRNCQNSNSNGFYCYRPLDS 294

>Contig20 gaMYB1
GGCACGAGACAGTTTCTCTTTTTGCTCATTTCCATCATCCCCATGTCCC
CTCTCCATGTTTTCTAAATCTCTCCATATACCTATAACACCGTTATTCT
TTCTCTATTCTACCTGATTTGATTTGATTTGATTTGTAAGTATGATGGGAC
GATCACCTTGTGTGAAAAGGCTCATACCAACAAAGGTGCCTGGACCAA
GAGGAAGATCAACGCCTCATCAACTACATCCGTGTCCATGGTGAAGGCTG
CTGGCGTTCCTCCCAAGCTGCTGGGCTGCTTAGATGTGGTAAGAGTT
GCAGATTAAGATGGATAAACTACTTGAGGCCTGATCTTAAGAGAGGAAAT
TTCAGTGAAGAAGAAGATGAGCTTATCATCAAGCTTCACAGTTTACTTGG
AAACAAATGGTCATTGATTGCTGGAAGATTACCAGGAAGGACAGATAATG
AGATAAAGAACTACTGGAACACACACATCAAAAGAAAGCTTATAAGCAGA
GGAATTGATCCACAACTCATCGTCCTCTCAATCAAACGGCCATTACCAA
CACAGTCACAGGCCCCACCGAATTGGATTTAGAACTCGCCACATCCG
TTTCCAAATCCAGTTCCATCAAAAACCGTCTCTGGATTTCAATTACAAT
GAATTTCAATTCAGTCCCACACAGATTCCTTGAAGAACCCAACTGTAC
AGCCAGCACTGGCATGACTACAGATGAAGAACAACAAGACAGCTGCACA
AGAAGCAGCAATACGGTCCGAGCAATGGGCAAGACATAAATTTGGAGCTG
TCGATTGGGATTGTTTCAGCTGACTCATCTCGGGTATCAAGTGCCAACTC
GGCCGAGTCGAAACCAAGGTAGATAACAACAATTTCCAGTTTCTTGAAC
AAGCTATGGTGGCTAAGGCGGTATGTTTGTGTTGGCAATTAGGTTTTGGA
ACAAGTGAAATTTGTAGGAAGTGTCAAAATTCAAATTCAAATGGCTTCTA
TAGTTATTGTAGACCTTGGATTATAGGGTCATCTTTTCTTCTTCTTCTT
TCTGTTTTTAGGAGATAAATTAATTCTTAATTATTATCTTTTGCCTGCC
CACTGTAATCAAAGCAAATTAATAAACAAGACTCTTTTGTATTTGTTT
AATTTATAGGTTCAAAAAAAAAAAAAAAAAAAAA

GaMYB1_alignment_and_sequence.txt

>Contig20_Frame+3 GaMYB1

HETVSLFCSFPSSPCPLSMFF*ISPYTYNTVILSLFYLI*FDLIL*LMGRSPCCEKAHTN
KGAWTKEEDQRLINYIRVHGEWCWRS LPKAAGLLRCGKSCRLRWINYLRPDLKRGNFTEE
EDELI IKLHSL LGNKWSLIAGRLPGR TDNEIKNYWNTHIKRKLISRGIDPQTHRPLNQTA
ITNTVTGPTELD FRNSPTSVSKSSSIKNPSLDFNYNEFHFKSHTDSLEEPNCTASTGMTT
DEEQQEQLHKKQQYGPSNGQDINLELSIGIVSADSSRVSSANSAESKPKVDNNNFQFLEQ
AMVAKAVCLCWQLGFGTSEICRNCQNSNSNGFYSYCRPLDS*GHLFLLSFCF*EIN*FLI
IIFLPAHCNQSKI IKQRLFLICSIYRFKKKKKK

>GaMYB1 - putative from GhMYB1 alignment

MGRSPCCEKAHTNKGAWTKEEDQRLINYIRVHGEWCWRS LPKAAGLLRCGKSCRLRWINY
LRPDLKRGNFTEEEDELI IKLHSL LGNKWSLIAGRLPGR TDNEIKNYWNTHIKRKLISRG
IDPQTHRPLNQTAITNTVTGPTELD FRNSPTSVSKSSSIKNPSLDFNYNEFHFKSHTDSL
EEP NCTASTGMTTDEEQQEQLHKKQQYGPSNGQDINLELSIGIVSADSSRVSSANSAESK
PKVDNNNFQFLEQAMVAKAVCLCWQLGFGTSEICRNCQNSNSNGFYSYCRPLDS